

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,418

DATE: 12/10/2001

TIME: 11:17:34

Input Set : N:\Crf3\RULE60\09977418.txt

Output Set: N:\CRF3\12102001\I977418.raw

5 <110> APPLICANT: Shimkets et al  
9 <120> TITLE OF INVENTION: Novel polynucleotides and polypeptides encoded thereby  
13 <130> FILE REFERENCE: 15966-552  
17 <140> CURRENT APPLICATION NUMBER: 09/977,418  
19 <141> CURRENT FILING DATE: 2001-10-15  
23 <150> PRIOR APPLICATION NUMBER: 09/584,411  
25 <151> PRIOR FILING DATE: 2000-05-31  
28 <150> PRIOR APPLICATION NUMBER: USSN 60/189,810  
30 <151> PRIOR FILING DATE: 2000-03-16  
33 <150> PRIOR APPLICATION NUMBER: USSN 60/191,158  
35 <151> PRIOR FILING DATE: 2000-03-22  
38 <150> PRIOR APPLICATION NUMBER: USSN 60/193,086  
40 <151> PRIOR FILING DATE: 2000-03-30  
43 <150> PRIOR APPLICATION NUMBER: USSN 60/201,388  
45 <151> PRIOR FILING DATE: 2000-05-03  
49 <160> NUMBER OF SEQ ID NOS: 93  
53 <170> SOFTWARE: PatentIn Ver. 2.0  
57 <210> SEQ ID NO: 1  
59 <211> LENGTH: 836  
61 <212> TYPE: DNA  
63 <213> ORGANISM: Artificial Sequence  
67 <220> FEATURE:  
69 <221> NAME/KEY: CDS  
71 <222> LOCATION: (189)..(695)  
75 <220> FEATURE:  
77 <223> OTHER INFORMATION: Description of Artificial Sequence: chemically  
79 synthesized  
83 <400> SEQUENCE: 1  
85 cagagtccct acccttttggga gaactgcgct tctcttttcgg agggagtgtt cgccgccgcc 60  
89 gcggccgccca cctggagttt cttcagactc cagatttccc tgtcaaccac gaggagtcca 120  
93 gagaggaaac gcggagcgga gacaacagta cctgacgcct ctttcagccc gggatcgccc 180  
97 cagcaggg atg ggc gac aag atc tgg ctg ccc ttc ccc gtg ctc ctt ctg 230  
99 Met Gly Asp Lys Ile Trp Leu Pro Phe Pro Val Leu Leu Leu  
101 1 5 10  
105 gcc gct ctg ctt cgg gtg ctg ctg cct ggg gcg gcc ggc ttc aca cct 278  
107 Ala Ala Leu Leu Arg Val Leu Leu Pro Gly Ala Ala Gly Phe Thr Pro  
109 15 20 25 30  
113 tcc ctc gat agc gac ttc acc ttt acc ctt ccc gcc ggc cag aag gag 326  
115 Ser Leu Asp Ser Asp Phe Thr Phe Thr Leu Pro Ala Gly Gln Lys Glu  
117 35 40 45  
121 tgc ttc tac cag ccc atg ccc ctg aag gcc tcg ctg gag atc gag tac 374  
123 Cys Phe Tyr Gln Pro Met Pro Leu Lys Ala Ser Leu Glu Ile Glu Tyr  
125 50 55 60  
129 caa gtt tta gat gga gca gga tta gat att gat ttc cat ctt acc tct 422  
131 Gln Val Leu Asp Gly Ala Gly Leu Asp Ile Asp Phe His Leu Thr Ser  
133 65 70 75  
137 cca gaa ggc aaa acc tta gtt ttt gaa caa aga aaa tca gat gga gtt 470

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,418

DATE: 12/10/2001

TIME: 11:17:34

Input Set : N:\Crf3\RULE60\09977418.txt

Output Set: N:\CRF3\12102001\I977418.raw

```

139 Pro Glu Gly Lys Thr Leu Val Phe Glu Gln Arg Lys Ser Asp Gly Val
141      80                      85                      90
145 cac act gta gag act gaa gtt ggt gat tac atg ttc tgc ttt gac aat 518
147 His Thr Val Glu Thr Glu Val Gly Asp Tyr Met Phe Cys Phe Asp Asn
149 95                      100                      105                      110
153 aca ttc agc acc att tct gag aag gtg att ttc ttt gaa tta atc ctg 566
155 Thr Phe Ser Thr Ile Ser Glu Lys Val Ile Phe Phe Glu Leu Ile Leu
157                      115                      120                      125
161 gat aat atg gga gaa cag gca caa gaa caa gaa gat tgg aag aaa tat 614
163 Asp Asn Met Gly Glu Gln Ala Gln Glu Gln Glu Asp Trp Lys Lys Tyr
165                      130                      135                      140
169 att act ggc aca gat ata ttg gat atg aaa ctg gaa gac atc ctg gac 662
171 Ile Thr Gly Thr Asp Ile Leu Asp Met Lys Leu Glu Asp Ile Leu Asp
173                      145                      150                      155
177 ctg ccc ggg cgg ccg ctc gag ccc tat agt gag taagtctgga ggcccggggcg 715
179 Leu Pro Gly Arg Pro Leu Glu Pro Tyr Ser Glu
181      160                      165
185 gccgctcctg cagtagggta ccgagctcgt cgacgcatgc tgatctagat cttaattaac 775
189 acgtggtgcc aagctttgga agactcagct tttgttcctt ttagtgaggg ttaatttcga 835
193 g 836
199 <210> SEQ ID NO: 2
201 <211> LENGTH: 169
203 <212> TYPE: PRT
205 <213> ORGANISM: Artificial Sequence
209 <220> FEATURE:
211 <221> NAME/KEY: variant
213 <222> LOCATION: (1)..(169)
215 <223> OTHER INFORMATION: where Xaa can any amino acid
219 <400> SEQUENCE: 2
221 Met Gly Asp Lys Ile Trp Leu Pro Phe Pro Val Leu Leu Leu Ala Ala
223 1                      5                      10                      15
227 Leu Leu Arg Val Leu Leu Pro Gly Ala Ala Gly Phe Thr Pro Ser Leu
229                      20                      25                      30
233 Asp Ser Asp Phe Thr Phe Thr Leu Pro Ala Gly Gln Lys Glu Cys Phe
235                      35                      40                      45
239 Tyr Gln Pro Met Pro Leu Lys Ala Ser Leu Glu Ile Glu Tyr Gln Val
241      50                      55                      60
245 Leu Asp Gly Ala Gly Leu Asp Ile Asp Phe His Leu Thr Ser Pro Glu
247 65                      70                      75                      80
251 Gly Lys Thr Leu Val Phe Glu Gln Arg Lys Ser Asp Gly Val His Thr
253                      85                      90                      95
257 Val Glu Thr Glu Val Gly Asp Tyr Met Phe Cys Phe Asp Asn Thr Phe
259                      100                      105                      110
263 Ser Thr Ile Ser Glu Lys Val Ile Phe Phe Glu Leu Ile Leu Asp Asn
265                      115                      120                      125
269 Met Gly Glu Gln Ala Gln Glu Gln Glu Asp Trp Lys Lys Tyr Ile Thr
271      130                      135                      140
275 Gly Thr Asp Ile Leu Asp Met Lys Leu Glu Asp Ile Leu Asp Leu Pro
277 145                      150                      155                      160

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,418

DATE: 12/10/2001

TIME: 11:17:34

Input Set : N:\Crf3\RULE60\09977418.txt

Output Set: N:\CRF3\12102001\I977418.raw

```

281 Gly Arg Pro Leu Glu Pro Tyr Ser Glu
283                               165
289 <210> SEQ ID NO: 3
291 <211> LENGTH: 2342
293 <212> TYPE: DNA
295 <213> ORGANISM: Homo sapiens
299 <220> FEATURE:
301 <221> NAME/KEY: CDS
303 <222> LOCATION: (110)..(1750)
307 <220> FEATURE:
309 <221> NAME/KEY: variation
311 <222> LOCATION: (1)..(711)
313 <223> OTHER INFORMATION: where n can be any nucleotide
317 <400> SEQUENCE: 3
319 agacccgctg agctgctagc ccgccggcca gcgagtgaga ggctggacag actgtggagc 60
323 cgacagactg aaggacagcg gcaccgccag acggccagaa agttccgcc atg agc tgg 118
325                               Met Ser Trp
327                               1
331 ggc acg gag ctg tgg gat cag ttc gac agc tta gac aag cat aca caa 166
333 Gly Thr Glu Leu Trp Asp Gln Phe Asp Ser Leu Asp Lys His Thr Gln
335      5                               10                               15
339 tgg gga att gac ttc ttg gaa aga tat gcc aaa ttt gtt aaa gag agg 214
341 Trp Gly Ile Asp Phe Leu Glu Arg Tyr Ala Lys Phe Val Lys Glu Arg
343 20                               25                               30                               35
347 ata gaa att gaa cag aac tat gcg aaa caa ttg aga aat ctg gtt aag 262
349 Ile Glu Ile Glu Gln Asn Tyr Ala Lys Gln Leu Arg Asn Leu Val Lys
351      40                               45                               50
355 aag tac tgc ccc aaa cgt tca tcc aaa gat gaa gag cca cgg ttt acc 310
357 Lys Tyr Cys Pro Lys Arg Ser Ser Lys Asp Glu Glu Pro Arg Phe Thr
359      55                               60                               65
363 tcg tgt gta gcc ttt ttt aat atc ctt aat gag tta aat gac tat gca 358
365 Ser Cys Val Ala Phe Phe Asn Ile Leu Asn Glu Leu Asn Asp Tyr Ala
367      70                               75                               80
371 gga cag cga gaa gtt gta gca gaa gaa atg gcg cac aga gtg tat ggt 406
373 Gly Gln Arg Glu Val Val Ala Glu Glu Met Ala His Arg Val Tyr Gly
375      85                               90                               95
379 gaa tta atg aga cat gct cat gat ctg aaa act gaa aga aaa atg cat 454
381 Glu Leu Met Arg His Ala His Asp Leu Lys Thr Glu Arg Lys Met His
383 100                               105                               110                               115
387 ctg caa gaa ggc cga aaa gct cac caa tct ctt gcc atg tgc tgg aac 502
389 Leu Gln Glu Gly Arg Lys Ala His Gln Ser Leu Ala Met Cys Trp Asn
391      120                               125                               130
395 cag atg gat aat agt aaa aag aag ttt gaa aga gaa tgt aga gag gca 550
397 Gln Met Asp Asn Ser Lys Lys Lys Phe Glu Arg Glu Cys Arg Glu Ala
399      135                               140                               145
403 gaa aag gcc cac cag agt tat gaa aga ttg gat aat gat act aat gca 598
405 Glu Lys Ala His Gln Ser Tyr Glu Arg Leu Asp Asn Asp Thr Asn Ala
407      150                               155                               160
411 acc aag gca gat gtt gaa aat gcc aaa cag cag ttg aat ctg cgt acg 646

```

## RAW SEQUENCE LISTING

DATE: 12/10/2001

PATENT APPLICATION: US/09/977,418

TIME: 11:17:34

Input Set : N:\Crf3\RULE60\09977418.txt

Output Set: N:\CRF3\12102001\I977418.raw

413	Thr	Lys	Ala	Asp	Val	Glu	Asn	Ala	Lys	Gln	Gln	Leu	Asn	Leu	Arg	Thr	
415		165					170					175					
419	cat	atg	gcc	gat	gaa	aat	aaa	aat	gca	tat	gct	gca	caa	tta	caa	aac	694
421	His	Met	Ala	Asp	Glu	Asn	Lys	Asn	Ala	Tyr	Ala	Ala	Gln	Leu	Gln	Asn	
423	180					185					190					195	
427	ttt	aat	gga	gaa	caa	cat	aaa	cat	ttt	tat	gta	gtg	att	cct	cag	att	742
429	Phe	Asn	Gly	Glu	Gln	His	Lys	His	Phe	Tyr	Val	Val	Ile	Pro	Gln	Ile	
431					200					205					210		
435	tac	aag	caa	cta	caa	gaa	atg	gac	gaa	cga	agg	act	att	aaa	ctc	agt	790
437	Tyr	Lys	Gln	Leu	Gln	Glu	Met	Asp	Glu	Arg	Arg	Thr	Ile	Lys	Leu	Ser	
439			215						220					225			
443	gag	tgt	tac	aga	gga	ttt	gct	gac	tca	gaa	cgc	aaa	gtt	att	ccc	atc	838
445	Glu	Cys	Tyr	Arg	Gly	Phe	Ala	Asp	Ser	Glu	Arg	Lys	Val	Ile	Pro	Ile	
447		230						235					240				
451	att	tca	aaa	tgt	ttg	gaa	gga	atg	att	ctt	gca	gca	aaa	tca	gtt	gat	886
453	Ile	Ser	Lys	Cys	Leu	Glu	Gly	Met	Ile	Leu	Ala	Ala	Lys	Ser	Val	Asp	
455		245					250					255					
459	gaa	aga	aga	gac	tct	caa	atg	gtg	gta	gac	tcc	ttc	aaa	tct	ggt	ttt	934
461	Glu	Arg	Arg	Asp	Ser	Gln	Met	Val	Val	Asp	Ser	Phe	Lys	Ser	Gly	Phe	
463	260					265					270					275	
467	gaa	cct	cca	gga	gac	ttt	cca	ttt	gaa	gat	tac	agt	caa	cat	ata	tat	982
469	Glu	Pro	Pro	Gly	Asp	Phe	Pro	Phe	Glu	Asp	Tyr	Ser	Gln	His	Ile	Tyr	
471					280					285					290		
475	aga	acc	att	tct	gat	ggg	act	atc	agt	gca	tcc	aaa	cag	gag	agt	ggg	1030
477	Arg	Thr	Ile	Ser	Asp	Gly	Thr	Ile	Ser	Ala	Ser	Lys	Gln	Glu	Ser	Gly	
479				295					300					305			
483	aag	atg	gat	gcc	aaa	acc	cca	gta	gga	aag	gcc	aag	ggc	aaa	ttg	tgg	1078
485	Lys	Met	Asp	Ala	Lys	Thr	Pro	Val	Gly	Lys	Ala	Lys	Gly	Lys	Leu	Trp	
487		310						315					320				
491	ctc	ttt	gga	aag	aag	cca	aag	ggc	cca	gca	cta	gaa	gat	ttc	agt	cat	1126
493	Leu	Phe	Gly	Lys	Lys	Pro	Lys	Gly	Pro	Ala	Leu	Glu	Asp	Phe	Ser	His	
495		325					330					335					
499	ctg	cca	cca	gaa	cag	aga	cgt	aaa	aaa	cta	cag	cag	cgc	att	gat	gaa	1174
501	Leu	Pro	Pro	Glu	Gln	Arg	Arg	Lys	Lys	Leu	Gln	Gln	Arg	Ile	Asp	Glu	
503	340					345					350					355	
507	ctt	aac	aga	gaa	cta	cag	aaa	gaa	tca	gac	caa	aaa	gat	gca	ctc	aac	1222
509	Leu	Asn	Arg	Glu	Leu	Gln	Lys	Glu	Ser	Asp	Gln	Lys	Asp	Ala	Leu	Asn	
511					360					365					370		
515	aaa	atg	aaa	gat	gta	tat	gag	aag	gat	cca	caa	atg	ggg	gat	cca	ggg	1270
517	Lys	Met	Lys	Asp	Val	Tyr	Glu	Lys	Asp	Pro	Gln	Met	Gly	Asp	Pro	Gly	
519			375						380					385			
523	agt	ttg	cag	cct	aaa	tta	gca	gag	acc	atg	aat	aac	att	gac	cgc	cta	1318
525	Ser	Leu	Gln	Pro	Lys	Leu	Ala	Glu	Thr	Met	Asn	Asn	Ile	Asp	Arg	Leu	
527			390					395					400				
531	cga	atg	gaa	atc	cat	aag	aat	gag	gct	tgg	ctc	tct	gaa	gtc	gaa	ggc	1366
533	Arg	Met	Glu	Ile	His	Lys	Asn	Glu	Ala	Trp	Leu	Ser	Glu	Val	Glu	Gly	
535		405					410					415					
539	aaa	aca	ggt	ggg	aga	gga	gac	aga	aga	cat	agc	agt	gac	ata	aat	cat	1414
541	Lys	Thr	Gly	Gly	Arg	Gly	Asp	Arg	Arg	His	Ser	Ser	Asp	Ile	Asn	His	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/977,418

DATE: 12/10/2001

TIME: 11:17:34

Input Set : N:\Crif3\RULE60\09977418.txt

Output Set: N:\CRF3\12102001\I977418.raw

```

543 420          425          430          435
547 ctt gta aca cag gga cga gaa agt cct gag gga agt tac act gat gat 1462
549 Leu Val Thr Gln Gly Arg Glu Ser Pro Glu Gly Ser Tyr Thr Asp Asp
551          440          445          450
555 gca aac cag gaa gtc cgt ggg cca ccc cag cag cat ggt cac cac aat 1510
557 Ala Asn Gln Glu Val Arg Gly Pro Pro Gln Gln His Gly His His Asn
559          455          460          465
563 gag ttt gat gat gaa ttt gag gat gat gat ccc ttg cct gct att gga 1558
565 Glu Phe Asp Asp Glu Phe Glu Asp Asp Asp Pro Leu Pro Ala Ile Gly
567          470          475          480
571 cac tgc aaa gct atc tac cct ttt gat gga cat aat gaa ggt act cta 1606
573 His Cys Lys Ala Ile Tyr Pro Phe Asp Gly His Asn Glu Gly Thr Leu
575          485          490          495
579 gca atg aaa gaa ggt gaa gtt ctc tac att ata gag gag gac aaa ggt 1654
581 Ala Met Lys Glu Gly Glu Val Leu Tyr Ile Ile Glu Glu Asp Lys Gly
583 500          505          510          515
587 gac gga tgg aca aga gct cgg aga cag aac ggt gaa gaa ggc tac gtt 1702
589 Asp Gly Trp Thr Arg Ala Arg Arg Gln Asn Gly Glu Glu Gly Tyr Val
591          520          525          530
595 ccc acg tca tac ata gat gta act cta gag aaa aac agt aaa ggt tcc 1750
597 Pro Thr Ser Tyr Ile Asp Val Thr Leu Glu Lys Asn Ser Lys Gly Ser
599          535          540          545
603 tgaagagggt ttctgaggaa atgggcaaga tggtgaagga gggtacatgc agctgctttt 1810
607 gggggagggt attagagttg tcaggctcaa agagagtgag agaagcaagt tgcattgagt 1870
611 catgcagaca tgattttttt ttactaact tcattagcat ttccatacat tgtttttaaa 1930
615 aatcataata ccaaccctta agttcctagt tcacagttat tcccacaaaa gaaaaagcca 1990
619 acaatagtgt accatttttc tattttatct tattgctgtc taatcaataa agaatgcaga 2050
623 ctgtgcaaaa aatgtgtctt acatttagct gtcccaacag gattgtcttc cctcccagct 2110
627 ctggttttaa ttggctttta gaccactat ctgtcagatc cttgccatct gtcagtgtct 2170
W--> 631 gcctgcgcca cctcgtgct tgcctaacat cctgttgcat gtctagcgtg attgagcnag 2230
W--> 635 attttcaggc atgtctttag aatcccctgg tncgtgtcaaa gcctgggtttg gtttacattg 2290
W--> 639 gtngtgcaat cncctttgtca acatctccag cactatngtt cctctctagg tn 2342
645 <210> SEQ ID NO: 4
647 <211> LENGTH: 547
649 <212> TYPE: PRT
651 <213> ORGANISM: Homo sapiens
655 <400> SEQUENCE: 4
657 Met Ser Trp Gly Thr Glu Leu Trp Asp Gln Phe Asp Ser Leu Asp Lys
659 1 5 10 15
663 His Thr Gln Trp Gly Ile Asp Phe Leu Glu Arg Tyr Ala Lys Phe Val
665 20 25 30
669 Lys Glu Arg Ile Glu Ile Glu Gln Asn Tyr Ala Lys Gln Leu Arg Asn
671 35 40 45
675 Leu Val Lys Lys Tyr Cys Pro Lys Arg Ser Ser Lys Asp Glu Glu Pro
677 50 55 60
681 Arg Phe Thr Ser Cys Val Ala Phe Phe Asn Ile Leu Asn Glu Leu Asn
683 65 70 75 80
687 Asp Tyr Ala Gly Gln Arg Glu Val Val Ala Glu Glu Met Ala His Arg
689 85 90 95

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/977,418

DATE: 12/10/2001

TIME: 11:17:35

Input Set : N:\Crf3\RULE60\09977418.txt

Output Set: N:\CRF3\12102001\I977418.raw

L:631 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:635 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:639 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:1159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:1161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:1215 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:1243 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:1247 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:1297 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8  
L:1375 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9  
L:1689 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11  
L:2295 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15  
L:2299 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15  
L:2303 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15  
L:3859 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21  
L:5903 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29  
L:5905 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29  
L:5987 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30  
L:9117 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45